



Published in final edited form as:

*Qual Life Res.* 2015 May ; 24(5): 1207–1216. doi:10.1007/s11136-014-0853-y.

## Identifying Changes in Scores on the EORTC-QLQ-C30 Representing a Change in Patients' Supportive Care Needs

Claire F. Snyder, PhD<sup>1,2,3</sup>, Amanda L. Blackford, ScM<sup>3</sup>, Jonathan Sussman, MD, MSc<sup>4</sup>, Daryl Bainbridge, PhD<sup>4</sup>, Doris Howell, RN, PhD<sup>5,6</sup>, Hsien Y. Seow, PhD<sup>4</sup>, Michael Carducci, MD<sup>3</sup>, and Albert W. Wu, MD, MPH<sup>1,2</sup>

<sup>1</sup>Division of General Internal Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA

<sup>2</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

<sup>3</sup>Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD, USA

<sup>4</sup>McMaster University, Hamilton, ON, Canada

<sup>5</sup>University Health Network/Princess Margaret Hospital, Toronto, ON, Canada

<sup>6</sup>University of Toronto, Toronto, ON, Canada

### Abstract

**Purpose**—Using health-related quality-of-life measures for patient management requires knowing what changes in scores require clinical attention. We estimated changes on the European Organization for Research and Treatment of Cancer Quality-of-Life-Questionnaire-Core-30 (EORTC-QLQ-C30) representing important changes by comparing to patient-reported changes in supportive care needs.

**Methods**—This secondary analysis used data from 193 newly-diagnosed cancer patients (63% breast, 37% colorectal; mean age 60 years; 20% male) from 28 Canadian surgical practices. Participants completed the Supportive Care Needs Survey-Short Form-34 (SCNS-SF34) and EORTC-QLQ-C30 at baseline, 3 weeks, and 8 weeks. We calculated mean changes in EORTC-QLQ-C30 scores associated with improvement, worsening, and no-change in supportive care needs based on the SCNS-SF34. Mean changes in the EORTC-QLQ-C30 scores associated with the SCNS-SF34 improved and worsened categories were used to estimate clinically important changes, and the 'no change' category to estimate insignificant changes.

**Results**—EORTC-QLQ-C30 score changes ranged from 6 to 32 points for patients reporting improved supportive care needs; statistically significant changes were 10-32 points. EORTCQLQ-C30 score changes ranged from 21-point worsening to 21-point improvement for patients reporting worsening supportive care needs; statistically significant changes were 9-21 points in the hypothesized direction and a 21-point statistically significant change in the opposite direction. EORTC-QLQ-C30 score changes ranged from a 1-point worsening to 16-point improvement for patients reporting stable supportive care needs.

---

**Corresponding Author:** Claire Snyder, PhD, Associate Professor of Medicine, 624 N. Broadway, Room 649, Baltimore, MD 21205, 443/287-5469, 410/955-0470 (FAX), csnyder@jhu.edu.

**CONFLICT OF INTEREST STATEMENT** The authors report no conflict of interest.

**Conclusion**—These data suggest 10-point EORTC-QLQ-C30 score changes represent changes in supportive care needs. When using the EORTC-QLQ-C30 in clinical practice, scores changing 10 points should be highlighted for clinical attention.

### Keywords

EORTC-QLQ-C30; supportive care needs; patient-reported outcomes; clinical practice; cancer; clinically important differences

---

## INTRODUCTION

One of the important applications of patient-reported outcome (PRO) measures is in routine clinical practice to inform patient management [1-12]. For this purpose, patients complete PROs and the resulting scores are provided to clinicians who may then use them to help identify and address issues in patients' functioning and well-being. However, this requires an understanding of what scores require clinical attention, including those that are poor in absolute terms and those representing an important worsening from a previous assessment. Scores that are poor in absolute terms are identified using a cut-off threshold at a single timepoint (e.g., all fatigue scores >30); score changes are determined by differences in scores at two timepoints (e.g., a fatigue score that worsened from 5-points at Time 1 to 20-points at Time 2 – a change of 15 points, which may be important even though it does not reach the absolute threshold of 30).

PRO scores representing potential problems, either in absolute terms or an important change, can be brought to the attention of clinicians to help inform their patient management. Notably, patients also frequently are able to access their PRO results and assess their progress [13]. Research has demonstrated that highlighting in some way the PRO scores on the results reports that may require clinical attention can help clinicians and patients apply the PRO results to improving the patient's care [14]. However, for many PRO measures, the absolute cut-off scores and important change scores are unknown. Research is needed to identify absolute cutoff scores and important score changes to inform interpretation of the PRO score reports, thereby improving the value of using PROs in clinical practice.

In previous research, we used a needs assessment to identify absolute cut-off scores associated with patient-reported unmet needs [15-16]. We first conducted an analysis using data from 117 breast, prostate, and lung cancer patients in the United States who completed both the European Organization for Research and Treatment of Cancer Quality-of-Life-Questionnaire-Core-30 (EORTC-QLQ-C30) and the Supportive Care Needs Survey-Short Form-34 (SCNS-SF34) at a single point in time [15]. These data were from a study that examined the health-related quality of life (HRQOL) and supportive care needs of cancer patients, and those two measures were selected based on preliminary research that indicated they were most relevant for clinical practice applications [17]. The study data were later used for a secondary data analysis to identify cut-off scores on the EORTC-QLQ-C30 associated with patient-reported unmet needs. Specifically, we examined the content of the SCNS-SF-34 to identify items potentially associated with EORTC-QLQ-C30 domains and

tested these relationships using receiver-operating characteristic (ROC) curve analysis. There were six EORTC-QLQ-C30 domains (physical function, role function, emotional function, pain, fatigue, global health/QOL) associated with SCNS-SF34 items as established by areas under the ROC curve (AUC) = 0.70. For these six domains, we identified cut-off scores on the EORTC-QLQ-C30 associated with unmet needs based on the SCNS-SF34 with sensitivity = 0.85 and specificity = 0.53. We later replicated the analyses in a sample of 408 Japanese breast cancer patients who completed the Japanese versions of the EORTC-QLQ-C30 and SCNS-SF34 [16]. The results of the Japanese data analysis were substantially similar to those from the original study, in terms of the EORTC-QLQ-C30 domains significantly associated with SCNS-SF34 items, and the sensitivity and specificity associated with the various cut-offs. Together, these two analyses provide useful information regarding absolute cut-off scores on the EORTC-QLQ-C30 to serve as a threshold for identifying patients with unmet needs and potentially requiring clinical attention.

However, because the previous datasets were cross-sectional, it was not possible to explore the changes in EORTC-QLQ-C30 scores that are associated with changes in patients' supportive care needs. In this analysis, we used data from a longitudinal study in which cancer patients completed the EORTC-QLQ-C30 and SCNS-SF34 at multiple time points. The purpose was to estimate score changes on the EORTC-QLQ-C30 representing changes in supportive care needs, and thus requiring clinical attention. Specifically, mean changes in the EORTC-QLQ-C30 scores associated with the SCNS-SF34 improved and worsened categories were used to estimate clinically important changes, and the 'no change' category was used to indicate insignificant changes. The results of this analysis were intended to facilitate the use of the EORTC-QLQ-C30 in clinical practice by identifying score changes that should be highlighted for clinical attention.

## METHODS

### Research Design and Data Source

This was a secondary analysis using data collected from a cluster randomized controlled trial (RCT) evaluating a community-based nursing-lead coordination of care intervention [18]. The original RCT was conducted in Toronto, Canada, and recruited newly diagnosed breast and colorectal cancer patients within 7 days of their surgery at 28 participating surgical clinics. Eligibility criteria included no previous or concomitant malignancies (except non-melanoma skin cancer or carcinoma in situ of the cervix), legally able to provide informed consent, 18 years of age or older, able to speak and read English, and residing in Toronto, ON.

Trained interviewers collected patients' QLQ-C30 and SCNS-SF34 questionnaires via telephone at baseline (2-7 days post-discharge from surgery), 3 weeks (2-3 weeks post-baseline), and 8 weeks (8-10 weeks post-baseline), along with other PRO measures. The data collection timepoints were based on the intervention being evaluated in the RCT and the estimated care trajectories for patients following cancer surgery. Specifically, the intent was to capture patient PROs following recovery (4 weeks after discharge) but before their formal entry into the cancer care system to begin treatment (10 weeks). Respondents were provided with a copy of the instrument to follow during the telephone interview. The

interviewer read the instructions on the survey, explained the response options, and repeated the response options as necessary during the interview. The original study found no significant effect of the intervention on PRO scores.

The EORTC-QLQ-C30 is a multi-dimensional HRQOL measure designed for use in cancer patients [19]. It includes five functional measures (physical, role, emotional, social, cognitive), eight symptoms (fatigue, pain, nausea/vomiting, appetite loss, constipation, diarrhea, insomnia, dyspnea), as well as global health/QOL and financial impact. Most items use a 4-item scale from 'not at all' to 'very much' and a one-week recall period. Raw scores are transformed to a 0-100 scale, with higher scores representing better functioning/QOL and greater symptom burden. The QLQ-C30 is used widely as both an outcome measure in clinical studies and a PRO in clinical practice [3, 5, 10, 20], and studies that have specifically investigated the appropriateness of questionnaires for clinical practice have supported its use [17, 21].

The SCNS-SF34 addresses unmet needs in five domains: physical and daily living, psychological, health system and information, patient care and support, and sexual [22-23]. For each item, respondents use a five-point scale: 1=not applicable (meaning they did not experience the issue), 2=satisfied (meaning the issue applies to them but is being adequately addressed), 3=low unmet need, 4=moderate unmet need, and 5=high unmet need. The instructions provide detailed explanations regarding each response category and explicitly label responses of "not applicable" and "satisfied" as "no need" and low, moderate, or high as "some need." The instructions and questionnaire can be downloaded from [24]. The recall period used in the RCT was 'since surgery' for baseline, or since the last survey for the week 3 and week 8 assessments.

The original study was approved by Hamilton Health Sciences Research Ethics Board, and the Johns Hopkins team was provided with a de-identified dataset. Thus, the Johns Hopkins School of Medicine Institutional Review Board deemed the current analysis as exempt.

## Analyses

The sample demographics were summarized using descriptive statistics. As a preliminary step to our main analysis, we first used the baseline data from the Canadian sample to replicate the analyses conducted in the original cross-sectional US and Japanese samples examining absolute cut-off scores [15-16]. This preliminary analysis was conducted to confirm that the associations between SCNS-SF34 items and EORTC-QLQ-C30 absolute cut-off scores previously established [15-16] were also present in this Canadian sample, thereby supporting our approach of using analogous methods to examine changes in scores over time – the novel aspect of the present analysis. The details of the absolute cut-off score analyses have been described [15-16]. Briefly, we calculated the AUC for each QLQ-C30 domain and potentially related SCNS-SF34 items and domains. Based on our previous analyses [15-16], we hypothesized that these six QLQ-C30 domains would be associated with these SCNS-SF34 items with AUC > 0.70: physical function with 'work around the home'; role function with 'work around the home'; emotional function with 'feelings of sadness', global health/QOL with 'feeling unwell a lot of the time'; pain with 'pain'; and fatigue with 'lack of energy/tiredness'. We calculated the associated sensitivity and

specificity for various cut-offs and compared them qualitatively with our previous findings [15-16].

We then proceeded with the analysis exploring changes in QLQ-C30 scores associated with changes in patient-reported supportive care needs, as measured by the SCNS-SF34. We first calculated the number of observations for each potential change in SCNS-SF34 responses for the relevant items. For example, observations that were 1=not applicable at baseline and 3=low unmet need at week 3 were counted as worsening because the patient went from no unmet need to some unmet need. As another example, observations that were 4=moderate unmet need at baseline and 2=satisfied at week 8 were counted as improvement because the patient went from some unmet need to no unmet need. This was done for changes between each pair of timepoints separately: baseline vs. week 3, baseline vs. week 8, and week 3 vs. week 8.

Figure 1 depicts changes on SCNS-SF34 items categorized as improvement, worsening, or unchanged. Because of sparse data for categories 3=low unmet need, 4=moderate unmet need, and 5=high unmet need, those three categories were combined into a single category “some unmet need.” Because of this merging across categories, changes within “some unmet need” (e.g., response of 5=high need at Time 1 and 3=low need at Time 2) were not considered in this analysis because they did not represent a change between “no need” and “some need,” nor did they represent no-change. Therefore, responses of 3=low need, 4=moderate need, or 5=high need at Time 1 and responses of 1=not applicable or 2=satisfied at Time 2 were categorized as improvement. The opposite changes were categorized as worsening (responses of 1 or 2 at Time 1 and 3, 4, or 5 at Time 2). The unchanged category included responses of 1=not applicable at both Time 1 and Time 2, and 2=satisfied at both Time 1 and Time 2. We combined the results across all three timepoints (e.g., observations of changes using baseline as Time 1 and week 3 as Time 2 were combined with observations of changes using week 3 as Time 1 and week 8 as Time 2). For each category (improved, worsening, unchanged), we calculated the mean changes in QLQ-C30 domain scores using intercept-only generalized estimating equation (GEE) linear regression models, to account for correlation among changes from the same patient. Effect sizes were calculated as the model estimate divided by the standard error (equivalent to the z-score). We did not adjust for multiple comparisons, and include  $p < 0.05$  to represent statistical significance for descriptive purposes only.

## RESULTS

A total of 193 patients participated in the RCT. The mean age was 60 years (range: 22 to 88), and 20% were male. The majority (63%) had breast cancer, and the remainder (37%) colorectal cancer. Over half (57%) had a college degree, and 62% were married. Of the 193 patients who completed the baseline assessment, 186 (96%) completed the week 3 assessment, and 179 (93%) completed the week 8 assessment.

The preliminary analyses comparing the results of the Canadian baseline data to the previous results regarding absolute cut-off scores in the US and Japanese sample produced generally similar results, though there were some differences of note. Three of the six QLQ-C30

domains that were associated with SCNS-SF34 items in the original [15] and Japanese [16] analyses with AUC 0.70 met the 0.70 threshold in these Canadian data: emotional function, pain, and global health/QOL (Table 1). Whereas for physical function and role function, the highest AUC was 0.69, and for fatigue, the highest AUC was 0.68. All of the QLQ-C30 domains that had AUCs<0.70 in our previous analyses [15-16] had AUCs<0.70 in this analysis.

Second, the SCNS-SF34 items found to have the strongest association with the QLQ-C30 domains were not always the same in the Canadian data as they had been in both the US and Japanese data, though most of these differences were minor. The association with physical function was higher for the SCNS-SF34 item 'feeling unwell a lot of the time' (AUC=0.69) than for 'work around the home' (AUC=0.67). The association with emotional function was higher for the SCNS-SF34 item 'feeling down or depressed' (AUC=0.79) than for 'feelings of sadness' (AUC=0.76). The difference for role function was somewhat greater, with the SCNS-SF34 item 'not being able to do the things you used to do' having an AUC=0.69 versus 'work around the home' with an AUC=0.58. Because of these differences, when calculating the sensitivity and specificity of the various QLQ-C30 cut-off scores, we investigated both the item with the strongest association from the previous analyses [15-16] and the item with the strongest association in the Canadian data (Table 2). The results were not substantially different using the SCNS-SF34 item with the strongest association in the Canadian data, compared to the SCNS-SF34 item with the strongest association from the US and Japanese analyses. Therefore, to maintain consistency across studies, we used the SCNS-SF34 item previously established as associated with the QLQ-C30 domain [15-16] in the remaining analyses.

Table 3 presents the changes in QLQ-C30 scores associated with improvement, worsening, and no-change on the associated SCNS-SF34 items. These change scores provide estimates of changes on these QLQ-C30 domains that represent a change in patient's unmet needs, and therefore provide interpretation guidance for future applications using the QLQ-C30 in clinical practice.

To summarize, 12 QLQ-C30 change scores were associated with improvements on the SCNS-SF34 [i.e., changes in the 6 QLQ-C30 domains for SCNS-SF34 items that went from (1) 'some unmet need' to 'not applicable' or (2) 'some unmet need' to 'satisfied']. Across these 12 changes, the number of observations ranged from 25 to 74. The QLQ-C30 mean improvement ranged from 6 to 32 points, with 11 changes reaching statistical significance (range 10-32 points). The absolute value of the effect sizes ranged from 1.70 to 7.90. In general, the changes in QLQ-C30 scores for patients who improved from 'some unmet need' to 'not applicable' were larger than for those who improved from 'some unmet need' to 'satisfied,' suggesting that there is a bigger change between not having the issue at all versus having the issue adequately addressed.

There were fewer observations for the 12 EORTC-QLQ-C30 change scores associated with SCNS-SF34 items categorized as worsened [i.e., changes in the 6 EORTC-QLQ-C30 domains for SCNS-SF34 items that went from (1) 'not applicable' to 'some unmet need' or (2) 'satisfied' to 'some unmet need']. Sample sizes ranged from 8 to 66. While the pattern of

EORTC-QLQC30 mean changes from ‘not applicable’ to ‘some unmet need’ was generally as hypothesized, the pattern from ‘satisfied’ to ‘some unmet need’ was less consistent, with 3 domains’ mean changes representing improvement rather than worsening. Overall, the EORTC-QLQ-C30 mean changes ranged from a 21-point worsening to a 21-point improvement. Four changes were statistically significant in the hypothesized direction (range 9-21 points), and one in the opposite direction (21 points). Effect sizes for associations in the hypothesized direction went as high as 3.70, but we also found effect sizes as high as 3.60 in the opposite direction.

The sample sizes for the 12 EORTC-QLQ-C30 change scores associated with SCNS-SF34 items that were unchanged were much larger, ranging from 49 to 217. Specifically, these 12 QLQ-C30 changes were calculated for SCNS-SF34 items that were (1) ‘not applicable’ at both timepoints or (2) ‘satisfied’ at both timepoints on the 6 EORTC-QLQ-C30 domains (physical, role, emotional, global health/QOL, pain, fatigue). Even though the magnitude of the mean QLQ-C30 change scores tended to be smaller for these categories, ranging from a 1-point worsening to 16-point improvement, 10 of the changes were statistically significant. Effect sizes ranged from 0.50 to 5.80 in absolute value.

## DISCUSSION

An important enabler of using PROs in clinical practice is guidance on how to interpret questionnaire scores, and in particular, identifying scores requiring attention [25-26]. In previous research, we were able to identify absolute cut-off scores on the QLQ-C30 associated with patient-reported unmet needs [15-16]. This information is useful, as a prior analysis demonstrated that the QLQ-C30 domains that were the poorest in absolute terms were most likely to predict the issues bothering patients the most [11]. It may also be useful to identify important changes in QLQ-C30 scores, and the results of the present analysis provide estimates of QLQ-C30 score changes associated with changes in patient’s needs. These estimates of meaningful change can be used in practice to highlight domains on the QLQ-C30 that may require clinical attention.

The approach taken in this analysis is similar to, but distinct from, analyses aiming to identify a minimally important difference (MID). This analysis provides estimates of clinically important changes representing changes in patients’ supportive care needs. It should be noted that the estimates for important changes were derived from group-level means, and that an individual patient’s perception of change may differ. Practically speaking, these estimates are intended to identify potential concerns for the clinician and patient to discuss, rather than to make a definitive diagnosis. Further, these estimates do not necessarily reflect the *smallest* difference that would be considered important and suggest the need for a change in management (i.e., the MID) [27], nor do they distinguish between minimal, moderate, and maximal important differences. Nevertheless, it is instructive to compare our findings to other research that has investigated interpretation of changes on the EORTC-QLQ-C30. Osoba et al. [28] and King [29] proposed categorizing changes on the QLQ-C30 of 5-10 points as small differences, 10-20 points as moderate differences, and greater than 20-point differences as large, with a 10-point change considered meaningful [30]. The findings of our study largely confirm these estimates.

For improvement, only one mean change was less than 10 points, seven mean changes were between 10 and 20 points, and four were higher than 20 points. The 11 changes that reached statistical significance ranged from 10 to 32 points. This suggests that the results from our novel approach of estimating changes in QLQ-C30 scores based on changes in supportive care needs are consistent with previous research aimed at estimating minimally important differences, at least for estimates of improvement.

The results of our analysis of score worsening are more difficult to interpret. Of the 12 SCNSSF34 change scores categorized as worsened, three mean changes in the hypothesized direction were less than 10 points, three were between 10 and 19 points, and two were 21 points. The four changes that were statistically significant in the hypothesized direction ranged from 9 to 21 points. However, four of the mean changes occurring in the physical function, role function, and pain domains were actually improvements, ranging from 0.5 to 21 points, with the 21-point change reaching statistical significance. These findings contrary to our hypotheses are consistent with the phenomenon of response shift.

Response shift is defined as a change in “internal standards, values, or conceptualizations” [31, p1115], leading to patients shifting how they would have responded retrospectively. Thus, a patient who initially rated their health as good, who then worsened but changed their valuation of health states (i.e., had a response shift), might now report their health as very good. In support of this explanation, Kvam et al. [32] found evidence of response shift on the QLQ-C30 in multiple myeloma patients. In the Kvam study, patients who worsened retrospectively reported better HRQOL at baseline for pain, fatigue, and physical function, but improving patients only demonstrated response shift on global QOL. The Kvam findings, in combination with the results from this study, suggest patients who are worsening may have shifted their standards or valuation of their previous HRQOL, leading to counter-intuitive results when compared with current reports on certain domains, particularly physical function and pain. What is interesting in our study is that both the SCNS-SF34 and QLQ-C30 were both self-reported measures collected simultaneously. Why would patients report worsening on the needs assessment but improvements on an HRQOL questionnaire? This question should be explored further to determine whether needs assessments may be less subject to response shift, and if so, why.

Another reason the results of our analysis would not be expected to establish ‘minimally’ important differences is because all of the unmet need categories (low, moderate, high) were collapsed into ‘some unmet need.’ There were not enough observations that changed between low and moderate and between moderate and high to examine changes of a single category. Further research using larger databases may be able to examine these smaller changes to provide a more refined estimation of the smallest difference representing a change in unmet needs.

The RCT that provided the data for this analysis assessed baseline after surgery, so we do not have scores for before surgery. However, because this analysis compared changes in EORTCQLQ-C30 scores with changes in SCNS-SF34 scores, regardless of when the changes occurred, it is not relevant whether the changes in scores occurred before or after surgery. We also combined the mean changes across all three time points (e.g., baseline to 3



weeks and baseline to 8 weeks). It may be instructive to examine the changes for each time point separately using larger datasets. Because the sample comprised 63% patients with breast cancer, the results here are most reflective of these patients, and there may be differences in the findings by cancer type. However, due to small sample sizes, we were unable to explore these differences fully.

Finally, the statistical significance of our results should be interpreted with caution. Due to differing sample sizes and associated power, some changes were small but statistically significant while others were large but statistically insignificant. Nevertheless, the magnitudes of the changes that we observed and effect sizes are generally informative. In summary, the results from this study add to the body of literature regarding the clinical significance of changes in QLQ-C30 scores that may require clinical attention. Our findings support changes of 10 points or greater as being clinically meaningful, particularly for improvements. The findings regarding worsening also suggest 10-points as being clinically meaningful, but due to the evidence of possible response shift, these findings should be interpreted with greater caution. Notably, the approach used here would be expected to provide larger than the 'minimal' changes representing a change in unmet needs, giving us confidence that these changes are important. Combining these results with previous research in this area contributes to the interpretability of the QLQ-C30, thereby facilitating its use in clinical practice to improve patient care.

## ACKNOWLEDGEMENT

This analysis was funded by the American Cancer Society (# MRS-08-011-01-CPPB). The original data collection was supported by the Canadian Health Services Research Foundation, Ontario Ministry of Health and Long-term Care. Drs. Snyder and Carducci are members of the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins (P30 CA 006973).

## ABBREVIATIONS

<b>AUC</b>	area under the curve
<b>EORTC-QLQ-C30</b>	European Organization for the Research and Treatment of Cancer Quality-of-Life-Questionnaire-Core-30
<b>GEE</b>	generalized estimating equation
<b>HRQOL</b>	health-related quality of life
<b>MID</b>	minimal important difference
<b>PRO</b>	patient-reported outcome
<b>RCT</b>	randomized controlled trial
<b>ROC</b>	receiver operating characteristic
<b>SCNS-SF34</b>	Supportive Care Needs Survey-Short Form-34

## REFERENCES

1. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. *The Lancet*. 2009; 374:369–70.

2. Greenhalgh J. The applications of PROs in clinical practice: What are they, do they work, and why? *Quality of Life Research*. 2009; 18:115–23. [PubMed: 19105048]
3. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *Journal of Clinical Oncology*. 2004; 22:714–24. [PubMed: 14966096]
4. Berry DL, Blumenstein BA, Halpenny B, et al. Enhancing patient-provider communication with the electronic self-report assessment for cancer: A randomized trial. *Journal of Clinical Oncology*. 2011; 29:1029–35. [PubMed: 21282548]
5. Detmar SB, Muller MJ, Schornagel JH, Wever LDV, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication. A randomized clinical trial. *JAMA*. 2002; 288:3027–34.
6. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: A literature review. *Journal of Evaluation in Clinical Practice*. 1999; 5:401–16. [PubMed: 10579704]
7. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: A structured review. *Journal of Evaluation in Clinical Practice*. 2006; 12:559–68. [PubMed: 16987118]
8. Haywood K, Marshall S, Fitzpatrick R. Patient participation in the consultation process: A structured review of intervention strategies. *Patient Education and Counseling*. 2006; 63:12–23. [PubMed: 16406464]
9. Cleeland CS, Wang XS, Shi Q, et al. Automated symptom alerts reduce postoperative symptom severity after cancer surgery: A randomized controlled clinical trial. *Journal of Clinical Oncology*. 2011; 29:994–1000. [PubMed: 21282546]
10. McLachlan S-A, Allenby A, Matthews J, et al. Randomized trial of coordinated psychosocial interventions based on patient self-assessments versus standard care to improve the psychosocial functioning of patients with cancer. *Journal of Clinical Oncology*. 2001; 19:4117–25. [PubMed: 11689579]
11. Snyder CF, Blackford AL, Aaronson NK, et al. Can patient-reported outcome measures identify cancer patients' most bothersome issues? *Journal of Clinical Oncology*. 2011; 29:1216–20. [PubMed: 21343558]
12. Snyder CF, Blackford AL, Wolff AC, et al. Feasibility and value of PatientViewpoint: a web system for patient-reported outcomes assessment in clinical practice. *Psycho-Oncology*. 2013; 22:895–901. [PubMed: 22544513]
13. Jensen RE, Snyder CF, Abernethy AP, et al. A Review of electronic patient reported outcomes systems used in cancer clinical care. *Journal of Oncology Practice*. 2013; 10:e215–22. [PubMed: 24301843]
14. Brundage, MD.; Bantug, ET.; Little, EA.; Smith, KA.; Snyder, CF. PRO Data Presentation Stakeholder Advisory Board. Which formats for communicating patient-reported outcomes (PROs) work best?. American Society of Clinical Oncology Annual Meeting; Chicago, Illinois. May 30-June 3, 2014;
15. Snyder CF, Blackford AL, Brahmer JR, et al. Needs assessments can identify scores on HRQOL questionnaires that represent problems for patients: an illustration with the Supportive Care Needs Survey and the QLQ-C30. *Quality of Life Research*. 2010; 19:837–45. [PubMed: 20340048]
16. Snyder CF, Blackford AL, Okuyama T, et al. Using the EORTC QLQC30 in clinical practice for patient management: Identifying scores requiring a clinician's attention. *Quality of Life Research*. 2013; 22:2685–91. [PubMed: 23532341]
17. Snyder CF, Dy SM, Hendricks DE, et al. Asking the right questions: Investigating needs assessments and health-related quality of life questionnaires for use in oncology clinical practice. *Supportive Care in Cancer*. 2007; 15:1075–1085. [PubMed: 17318591]
18. Sussman J, Howell D, Bainbridge D, et al. Results of a cluster randomized trial to evaluate a nursing lead supportive care intervention in newly diagnosed breast and colorectal cancer patients. *Journal of Clinical Oncology*. 2012; 30(suppl) Abstract 6035.

19. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*. 1993; 85:365–76. [PubMed: 8433390]
20. Lipscomb, J.; Gotay, CC.; Snyder, C., editors. *Outcomes Assessment in Cancer: Measures, Methods, and Applications*. Cambridge University Press; Cambridge: 2005.
21. Snyder CF, Herman JM, White SM, et al. When using patient-reported outcomes in clinical practice, the measure matters: A randomized controlled trial. *Journal of Oncology Practice*. 2014; 10:e299–306. [PubMed: 24986113]
22. Bonevski B, Sanson-Fisher RW, Girgis A, et al. Evaluation of an instrument to assess the needs of patients with cancer. *Cancer*. 2000; 88:217–25. [PubMed: 10618626]
23. Sanson-Fisher R, Girgis A, Boyes A, et al. The unmet supportive care needs of patients with cancer. *Cancer*. 2000; 88:226–37. [PubMed: 10618627]
24. National Cancer Institute. Grid-Enabled Measures Database: Supportive Care Needs Survey-Short Form (SCNS-SF34). Available at: <https://www.gem-beta.org/Public/MeasureDetail.aspx?mid=1592&cat=2>
25. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: A review of the options and considerations. *Quality of Life Research*. 2012; 21:1305–14. [PubMed: 22048932]
26. International Society for Quality of Life Research. User's Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice, Version: November 11, 2011. Available at: <http://www.isoqol.org/research/isoqol-publications>
27. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*. 1989; 10:407–15.
28. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*. 1998; 16:139–44. [PubMed: 9440735]
29. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Quality of Life Research*. 1996; 5:555–67. [PubMed: 8993101]
30. Maringwa JT, Quinten C, King M, et al. Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Supportive Care in Cancer*. 2011; 19:1753–60. [PubMed: 20886240]
31. Ahmed S, Schwartz C, Ring L, Sprangers MAG. Applications of health-related quality of life for guiding health care: Advances in response shift research. *Journal of Clinical Epidemiology*. 2009; 62:1115–7. [PubMed: 19595571]
32. Kvam AK, Wisloff F, Fayers PM. Minimal important differences and response shift in health-related quality of life; a longitudinal study in patients with multiple myeloma. *Health and Quality of Life Outcomes*. 2010; 8:79. [PubMed: 20678240]

a) Responses Categorized as **Improvement**: 3=low need, 4=moderate need, or 5=high need at Time 1 (open circles) and responses of 1=not applicable or 2=satisfied at Time 2 (shaded circles)

... what was your level of need for help with:	Not applicable	Satisfied	Low need	Moderate need	High need
	1	2	3	4	5

b) Responses Categorized as **Worsening**: 1=not applicable or 2=satisfied at Time 1 (open circles) and responses of 3=low need, 4=moderate need, or 5=high need at Time 2 (shaded circles)

... what was your level of need for help with:	Not applicable	Satisfied	Low need	Moderate need	High need
	1	2	3	4	5

c) Responses Categorized as **Unchanged**: 1=not applicable at both Time 1 (open circle) and Time 2 (shaded circle) and 2=satisfied at both Time 1 (open circle) and Time 2 (shaded circle)

... what was your level of need for help with:	Not applicable	Satisfied	Low need	Moderate need	High need
	1 1	2 2	3	4	5

**Figure 1.**  
Changes in SCNS-SF34 Scores Representing Improvement, Worsening, and No Change

**Table 1**

Hypothesized Relationship between QLQ-C30 and SCNS-SF34 Domains and Resulting Areas Under the Curve (AUC): Original, Japanese, and Canadian Analyses

QLQ-C30 Domain	SCNS-SF34 Domain/Item(s)	AUC		
		Original Analysis	Japanese Data	Canadian Data
<i>AUC 0.70 in previous studies [15-16]</i>				
Physical Function	Physical & Daily Living Needs (overall score and individual items)	0.69-0.81	0.69-0.74	0.59-0.69
Role Function	Work around the home Not being able to do the things you used to	0.71-0.73	0.70-0.70	0.58-0.69
Emotional Function	Psychological Needs (overall score and individual items)	0.56-0.74	0.61-0.75	0.67-0.79
Pain	Pain	0.78	0.74	0.75
Fatigue	Lack of energy/tiredness	0.74	0.75	0.68
Global Health /QOL	Feeling unwell a lot of the time	0.73	0.76	0.72
<i>AUC &lt;0.70 in previous studies [15-16]</i>				
Social Function	Not being able to do the things you used to	0.64	0.68	.65
Sleep	Lack of energy/tiredness Feeling unwell a lot of the time Being given information...about aspects of managing your illness and side-effects at home	0.41-0.51	0.39-0.55	0.47-0.56
Cognitive Function	Feeling unwell a lot of the time Being given information...about aspects of managing your illness and side-effects at home	0.54-0.60	0.53-0.63	0.48-0.51
Nausea/Vomiting		0.19-0.36	0.22-0.27	0.41-0.44
Dyspnea		0.37-0.48	0.32-0.48	0.22-0.23
Appetite Loss		0.47-0.49	0.32-0.49	0.35-0.48
Constipation		0.31-0.37	0.32-0.40	0.35-0.40
Diarrhea		0.34-0.34	0.18-0.21	0.18-0.19

**Table 2**

Sensitivity and Specificity of Various Cut-Off Scores: Original [15], Japanese [16], and Canadian Analyses

QLQ-C30 Domain	SCNS-SF34 Item	Cut-Off	Cohort	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Physical Function	<sup>1</sup> Work around the home <sup>2</sup> Feeling unwell a lot of the time	80	Original	0.65	0.83	0.55	0.89
			Japanese	0.40	0.92	0.63	0.82
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.77 0.79	0.56 0.51	0.52 0.34	0.80 0.88
		90	Original	0.85	0.58	0.39	0.92
			Japanese	0.85	0.65	0.45	0.93
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.95 0.91	0.28 0.23	0.45 0.28	0.89 0.89
Role Function	<sup>1</sup> Work around the home <sup>2</sup> Not being able to do the things you used to do	80	Original	0.69	0.79	0.50	0.89
			Japanese	0.69	0.79	0.52	0.88
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.91 0.94	0.26 0.32	0.43 0.54	0.82 0.87
		90	Original	0.85	0.69	0.46	.94
			Japanese	0.85	0.62	0.43	0.93
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.97 0.99	0.13 0.16	0.41 0.50	0.89 0.94
Emotional Function	<sup>1</sup> Feelings of sadness <sup>2</sup> Feeling down or depressed	90	Original	0.89	0.53	0.48	0.91
			Japanese	0.84	0.60	0.58	0.86
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.90 0.92	0.43 0.43	0.37 0.38	0.92 0.94
		100	Original	0.94	0.35	0.41	0.93
			Japanese	0.92	0.42	0.51	0.89
			Canadian <sup>1</sup> Canadian <sup>2</sup>	0.98 1.00	0.22 0.23	0.32 0.32	0.97 1.00
Global Health/QOL	Feeling unwell a lot of the time	70	Original	0.71	0.69	0.52	0.84
			Japanese	0.86	0.56	0.33	0.94
			Canadian	0.96	0.34	0.32	0.96
		80	Original	0.89	0.58	0.50	0.91
			Japanese	0.89	0.45	0.29	0.94
			Canadian	0.98	0.24	0.30	0.97
Pain	Pain	20	Original	0.66	0.84	0.64	0.85
			Japanese	0.70	0.81	0.62	0.86
			Canadian	0.90	0.46	0.32	0.95
		10	Original	0.91	0.66	0.54	0.95
			Japanese	0.93	0.54	0.47	0.94
			Canadian	0.95	0.26	0.26	0.95

QLQ-C30 Domain	SCNS-SF34 Item	Cut-Off	Cohort	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Fatigue	Lack of energy/ tiredness	30	Original	0.77	0.71	0.73	0.75
			Japanese	0.86	0.62	0.54	0.90
			Canadian	0.94	0.32	0.44	0.91
		20	Original	0.91	0.55	0.68	0.86
			Japanese	0.97	0.42	0.46	0.97
			Canadian	0.99	0.19	0.40	0.96

<sup>1</sup>Using same item as Original and Japanese data;

<sup>2</sup>Using Canadian item with highest AUC

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3** Mean QLQ-C30 Score Changes [Effect Size] (95% Confidence Intervals) Associated with Changes in Supportive Care Needs

	Physical Function*	Role Function*	Emotional Function*	Global Health/QOL*	Pain <sup>^</sup>	Fatigue <sup>^</sup>
<b>SCNS-34 Improved</b>						
Some Unmet Need to Not Applicable	n=50 17.3 [6.60] (12.2 to 22.4) p<0.001	n=50 32.3 [7.90] (24.3 to 40.4) p<0.001	n=25 16.7 [4.10] (8.8 to 24.6) p<0.001	n=34 20.1 [5.50] (12.9 to 27.3) p<0.001	n=36 -31.0 [-6.70] (-40.1 to -21.9) p<0.001	n=27 -25.9 [-5.80] (-34.8 to -17.1) p<0.001
Some Unmet Need to Satisfied	n=58 15.2 [5.50] (9.8 to 20.6) p<0.001	n=58 18.1 [4.20] (9.7 to 26.5) p<0.001	n=33 5.8 [1.70] (-0.8 to 12.4) p=0.10	n=35 9.8 [3.00] (3.3 to 16.2) p=0.006	n=42 -28.6 [-5.60] (-38.5 to -18.7) p<0.001	n=74 -19.0 [-6.90] (-24.4 to -13.6) p<0.001
<b>SCNS-34 Worsened</b>						
Not Applicable to Some Unmet Need	n=13 3.1 [0.50] (-10.2 to 16.4) p=0.66	n=13 -12.8 [-1.40] (-30.6 to 5.0) p=0.18	n=26 -21.2 [-3.70] (-32.3 to -10.0) p=0.001	n=18 -13.4 [-2.50] (-24.0 to -2.8) p=0.02	n=12 15.3 [2.10] (1.2 to 29.4) p=0.06	n=8 20.8 [3.70] (9.7 to 32.0) p=0.008
Satisfied to Some Unmet Need	n=41 5.4 [1.50] (-1.6 to 12.4) p=0.14	n=41 20.7 [3.60] (9.3 to 32.2) p<0.001	n=36 -9.3 [-2.60] (-16.2 to -2.4) p=0.01	n=39 -4.9 [-1.30] (-12.1 to 2.3) p=0.19	n=37 -0.5 [-0.10] (-8.3 to 7.4) p=0.91	n=66 1.5 [0.50] (-4.7 to 7.8) p=0.64
<b>SCNS-34 Unchanged</b>						
Not Applicable to Not Applicable	n=72 6.7 [4.80] (4.0 to 9.5) p<0.001	n=72 15.7 [4.50] (8.9 to 22.6) p<0.001	n=192 3.4 [3.60] (1.6 to 5.3) p<0.001	n=217 6.1 [4.90] (3.6 to 8.5) p<0.001	n=122 -2.7 [-2.00] (-5.4 to 0) p=0.05	n=52 -6.0 [-2.20] (-11.3 to -0.7) p=0.03
Satisfied to Satisfied	n=107 4.4 [2.70] (1.2 to 7.6) p=0.008	n=107 7.8 [2.50] (1.6 to 14.0) p=0.01	n=53 -1.1 [-0.50] (-5.1 to 2.9) p=0.60	n=49 6.1 [1.70] (-0.8 to 13.1) p=0.09	n=112 -12.9 [-5.80] (-17.3 to -8.6) p<0.001	n=125 -5.6 [-2.80] (-9.6 to -1.7) p=0.006

\* For these domains, positive changes reflect improvement in functioning and negative scores reflect worsening in functioning

<sup>^</sup> For these domains, positive changes reflect worsening in symptoms and negative scores reflect improvement in symptoms

**BOLD: p 0.05**

Gray shading: change in direction contrary to hypothesis